

Infrared Small Target Tracking Based on Target Spatial Distribution With Improved Kernelized Correlation Filtering

ZhiQiang Kou

Xinjiang University

Askar Hamdulla (✉ askar@xju.edu.cn)

Xinjiang University, Department of Information Science and Technology, Urumqi, 830002 CHN

Research Article

Keywords: correlation filter, Multi feature function, re-detection mechanism, infrared (IR) small target

Posted Date: January 5th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-135597/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Infrared small Target Tracking Based on Target Spatial Distribution with Improved kernelized Correlation Filtering

Zhiqiang Kou¹, Askar Hamdulla^{2*}

¹Xinjiang University, Department of Information Science and Technology, Urumqi, 830002 CHN

²Xinjiang University, Department of Information Science and Technology, Urumqi, 830002 CHN

*Corresponding author: Askar Hamdulla (askar@xju.edu.cn).

ABSTRACT

The application of correlation filtering in infrared small target tracking has been a mature research field. Traditional correlation filtering is to describe the target features by using a single feature, which can not solve the problem of target occlusion. Because of the fast moving speed and lack of re-detection mechanism, the target tracking will produce offset, which leads to the performance of the tracker to decline. In view of the above problems, a new multi feature re detection framework is proposed for long-term tracking of small targets. The feature selects multi feature weighting function, considers the importance of intensity feature to infrared target and different regions, calculates the gray distribution weighting function of the target, and combines the weighting function into the correlation filter. Before updating the template, to verify the reliability of target detection, the average peak correlation energy is used as the confidence of candidate region. When the target is completely occluded, the prediction result of Kalman filter is used as the optimal estimation of target position in the next frame. A large number of experimental results on different video sequences show that the tracking accuracy of this method is greatly improved compared with the baseline method.

Index Terms: correlation filter, Multi feature function, re-detection mechanism, infrared (IR) small target.

Introduction

Infrared target detection and tracking is of great significance because of its comprehensive military applications, such as video surveillance, infrared imaging precision guidance, visual monitoring and so on[1].The main challenge in these areas is that small infrared targets are fuzzy in space, and their size in the image is several pixels, which is easy to be confused with part of the noise[2]. For infrared small target detection, our team has created many algorithms, such as [3]. Du et al. Proposed an infrared small target detection algorithm based on human visual system. In [4], Wu proposed a space-time infrared small target detection method. In this paper, the infrared target tracking algorithm is mainly introduced. Infrared imaging tracking is the same as the traditional tracking, which calculates the position of the target in the next frame image sequence and tracks it. Next, we will introduce the current mainstream infrared small target tracking algorithm, and analyze their performance.

At present, there are several commonly used infrared small target tracking frame works. One is to use the difference between the target and the background to preprocess and enhance the small target, and then track the enhanced infrared small target. As shown in [5], Qian et al. Used SVD algorithm to estimate the background of infrared image and enhance weak and small targets. Finally, the response mapping of kernel classifier is used to estimate the target position.

Here's another way to use features directly to describe target features. Given the position of the target in the first frame, a frame is established as the candidate target with the largest correlation with the target in the first frame by correlation. As described in reference [6], a low rank sparse correlation filter based on mathematical processing error is proposed.

In this paper, a correlation filtering model is proposed for long-term tracking of small infrared targets, as shown in Fig. 1. Using the core technology of KCF algorithm, the computing time is reduced. Secondly, KCF algorithm uses a single feature, no re detection mechanism, occlusion processing effect is not good. In this method, GF multi feature function is selected to enhance the description of small target and a new re-detection mechanism is used to correct the tracking results, which effectively solves the problem of incomplete occlusion. When the target is completely occluded, Kalman filter is used to predict the target.

It is stated here that the proposed method is only applicable to the scene with complex background and severe occlusion, and is not suitable for the scene without occlusion and less background noise. When the above situation is the case, the baseline algorithm is adopted because of the fast calculation speed.

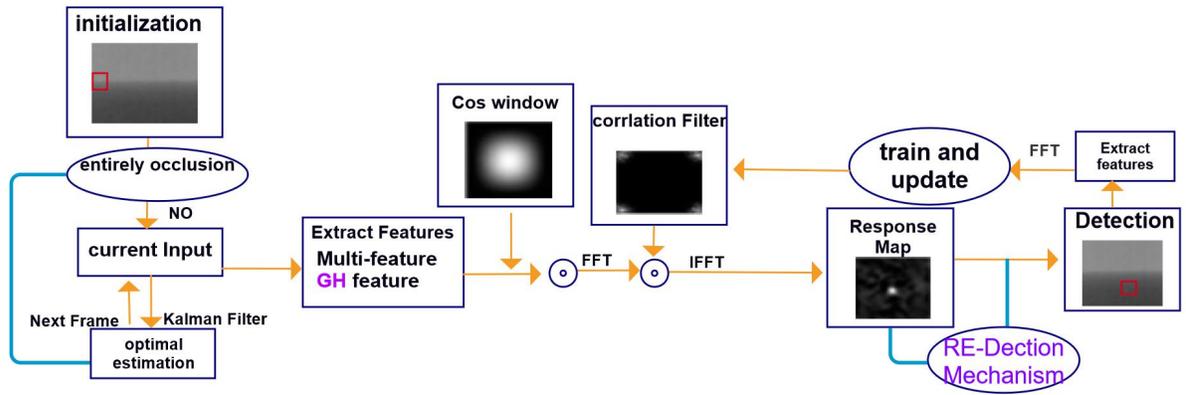


Figure 1 algorithm flow chart

Methods

- Tradition correlation filter (CF) tracker

The tracker trained by CF can be expressed as a minimum ridge regression problem [7], such as Formula 1. λ Indicates that the regular coefficient is used to prevent overfitting, $f(x_i) = \langle w, x_i \rangle + b$, $\langle \cdot \rangle$ means dot product.

$$\min_i \sum_i (f(x_i) - y_i)^2 + \lambda \|w\|^2 \quad (1)$$

The KCF algorithm uses several techniques to solve the above problems. The first is to select the cyclic matrix to increase the positive and negative samples in the target training. The second is to use the diagonalization of the cyclic matrix to greatly reduce the calculation time. The second is to use the kernel function to map the linear problem to the non-linear kernel space .

KCF algorithm is a good choice in the field of infrared small target tracking [5], because it is very fast, but the tracking effect is not good in complex background and occlusion background, because of its single feature and lack of re detection mechanism. We propose a new framework to deal with the above problems, and then we will introduce the proposed method.

- Multi-feature function

Next, we will analyze the characteristics of the target. When the target does not enter the cloud, the gradient change of the target is obvious, as shown in the second column of Figure 2. However, when the target enters the cloud, the gradient feature of the target is not particularly important, as shown in the first column of figure 2

When the target enters the cloud, the traditional KCF algorithm will lose the target, as shown in Figure 3. This is because a single gradient feature cannot uniquely represent the characteristics of small targets, which will lead to tracking failure.

Therefore, it is not enough to use hog features only. We need to add additional features to increase the representation of IR targets. For small infrared targets, intensity is a widely used feature because it is not affected by scaling and rotation, and it is highly sensitive to complex backgrounds.

Considering the importance of intensity characteristics, the intensity characteristics of the target cannot be well reflected in the KCF algorithm. So a new KCF algorithm with multiple feature functions is proposed. The definition of multi characteristic function is shown in formula 2.

$$F = \lambda F_{gary} + \alpha F_{HOG} \quad (2)$$

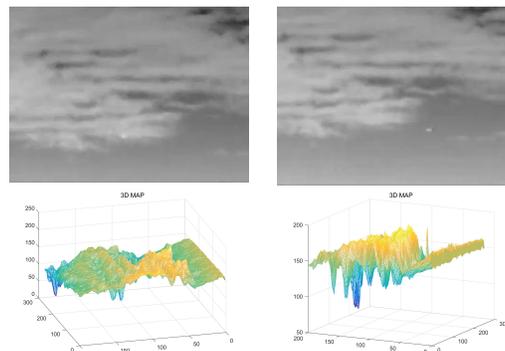


Figure 2 3D image under occlusion condition and 3D image without occlusion

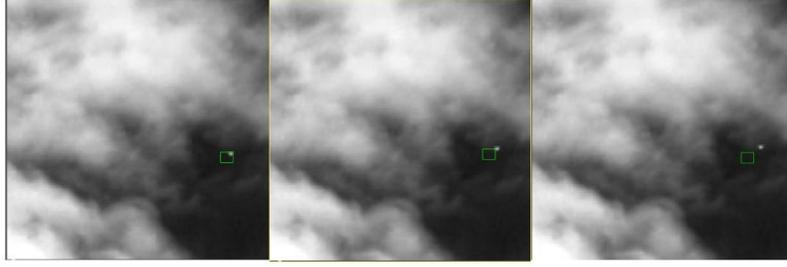


Figure 3 KCF algorithm fails to track when the target enters the cloud. (The green box represents the tracking result.)

F_{HOG} is the HOG feature of KCF algorithm, and F_{gray} represents the distribution feature of small targets. Where λ , and α represent the coefficients of gray feature and HOG feature. The specified coefficient is restricted to 1, that is, $\lambda + \alpha = 1$. The gradient descent method is adopted to find the optimal solution of the coefficient and update it online [8]. The gray features are described below.

The gray feature of image is represented by gray histogram, which represents the frequency of each gray level. The number of channels of gray histogram is 255, which is expressed as $q = \{q_u\}_{u=1}^L$.

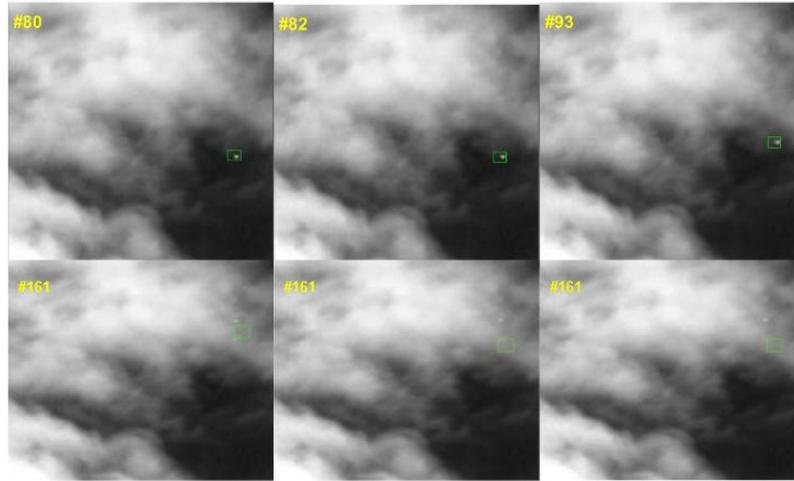


Figure 4 Tracking result map after adding gray feature and hog multi feature function.

After adding the multi feature function of gray features and hog features, the algorithm can achieve very good results when the cloud enters the cloud, but when the target enters the thick cloud layer, the tracking failure will occur again, as shown in Figure 4 (the first line is the tracking result when the target is in the thin cloud layer, and the second line is the failure tracking result when the target is in the thick cloud layer). This is because of the gray of the target the degree information is too close to the gray information of the background cloud, so the contrast between the target and the cloud can not be increased by using the gray information only.

For enhance the contrast between the target and the heavy cloud, we propose a gray distribution feature to enhance the weight information of the target, so as to enhance the effect of the target. In other words, the importance of pixels in each position is different. The pixels close to the target are important, while the pixels far away from the target are usually less important [9]. The pixels in the target area are weighted by formula 3, d is defined as the distance between the pixel and the target center.

$$y(d) = \begin{cases} 1 - d^2 & d < 1 \\ 0 & d \geq 1 \end{cases} \quad (3)$$

The intensity distribution histogram is defined as formula 4, where x_0 is the coordinate of the target center.

$$q_u(x_i) = \sum_{i=1}^W y\left(\left\|\frac{x_0 - x_i}{h}\right\|^2\right), u = 1, \dots, L \quad (4)$$

h is the size of the target area defined as $h = \sqrt{h_x^2 + h_y^2}$, W is the quantity of pixels in the target region, x_i is the position of the pixel. After adding the multi feature function of gray distribution feature and hog feature, the contrast pair between the target and thick cloud is added to deal with the tracking failure when the target enters the cloud. The experimental results after adding the new multi feature function are shown in Fig. 5.



Figure 5 Tracking results of algorithm with new multi feature function

- Re-detection model

When the KCF algorithm is updated with the new template, the accuracy of the algorithm is not considered. The fact is that once the target detection is inaccurate, it is difficult to correct it by the algorithm itself, which may cause tracking failure. So in this article, a re-inspection mechanism is added, it is used to control the template update, the re-detection model will be introduced next.

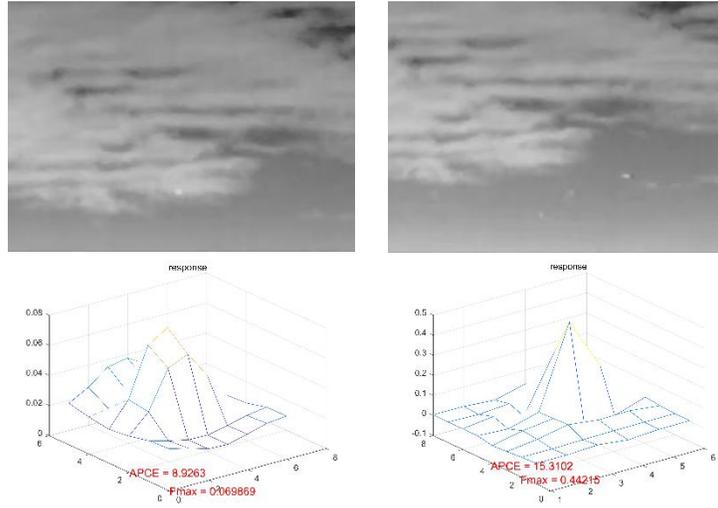


Figure 6 The difference map of APCE between occlusion and non occlusion, The first behavior is the original image with occlusion and the original image without occlusion, and the second behavior is the corresponding APCE.

During the target detection process, the tracking results are used by us to determine the necessity of model update [10]. We use the average peak correlation energy [11] to judge the response. If the target position corresponding to the maximum response value is not our target [11], it will be detected again. APCE represents the degree of fluctuation of the response graph and the confidence level of the detection target [11], which is defined as formula 5. In different situations, the target has different APCE. For example, when the target is occluded, the APCE will decrease significantly, as shown in Figure 6.

$$APCE = \frac{|F_{max} - F_{min}|^2}{\text{mean}_{\sum_{w,h}} (F_{w,h} - F_{min})^2} \quad (5)$$

Where F_{max} , F_{min} and $F_{w,h}$ represent the maximum response value, minimum response value and the unit in the w th row and h th column of the response graph. Adding the re-detection model, as shown in the following table. Where Mean_{APCE} and $\text{Mean}_{F_{max}}$ are defined as formula 6 and formula 7.

Table 1 Re-detection flow chart

Proposed Re-detection mechanism
INPUT: Initial target box (X,Y,W,H)
Output: Filter template after training
detection: Find F max and APCE according to the filter template after training
IF $F_{max} < \text{Mean}_{F_{MAX}}$ or $APCE < \text{Mean}_{APCE}$
Re-detection the image-Patch
Calculate the maximum response value FMAX in the candidate area
End
End
Get the target location
Update the model
Until end of mage sequence

$$\text{Mean}_{APCE} = \left(\frac{\sum_i^n APCE_1 + \dots + APCE_i}{i} \right) \quad (6)$$

$$\text{Mean}_{F_{\max}} = \left(\frac{\sum_i^n F_{\max_1} + \dots + F_{\max_i}}{i} \right) \quad (7)$$

- Kalman filter

Long-term tracking is the problem to be solved by this article. The processing method of occlusion and tracking loss is designed, namely Kalman filter trajectory prediction.

Kalman filter can correct the tracking coordinates of the tracker output. In case of complete occlusion or similar target interference, Kalman filter can predict the possible position of the next frame coordinates.

Results

In this section, we compare our algorithm with four baseline algorithms: KCF tracker, CSK tracker, CXT tracker and DFT tracker [8]-[11]. Among the six challenging infrared sequences, the sequence details are shown in Table 2. The evaluation of experimental speed is shown in Table 1.

The sequence data set contains the following attributes: tracking under thin clouds, tracking under thick clouds, tracking under full occlusion, tracking background clutter, fast moving target, low resolution of tracking sequence, low contrast between target and background. The specific information is shown in Table 1. The resolution of image sequence is 250×250 , 250×250 , 504×396 , 250×250 , 256×200 . The tracking results are shown in Fig6.

All experiments have been carried out in MATLAB r2016a on a computer with 16GB memory and 2.66-GHz Intel i7-920 processor.

Table 2 Image sequence details table

Sequence	Number	Target	Target detail	Background detail
Seq.1	168	aircraft	·Point target ·incomplete occlusion	·Thin clouds ·Changing background
Seq.2	30	aircraft	·Point target ·Heavy occlusion	·Heavy clouds ·Almost Keeping the same
Seq.3	429	Vehicle	·low contrast ·Point target	·Heavy noise ·background complex
Seq.4	381	Rocket	·Fast moving ·Illumination changing	·Camera moving ·Changing backgrounds
Seq.5	342	Ship	·low contrast	·backgrounds clutter
Seq.6	30	airplane	·Fast moving ·low contrast	·Low contrast between cloud and target

- Subjective experimental analysis

The first scene is cloud sky background. The similarity between the target and the background is very high, as shown in the first line of Figure 8. When the target just enters the cloud, the KCF algorithm fails to track. This is because the KCF algorithm has poor tracking robustness under the condition of target occlusion, and the DFT algorithm has already had the problem of tracking failure in the previous needle. CSK, CXT and our method can effectively track small targets.

When the target enters the thick cloud, as shown in the second line of Figure 8. KCF algorithm and DFT algorithm have failed to track before, while CXT algorithm has failed in 161 frames. This is because CXT algorithm uses a single feature and does not enhance the contrast between target and cloud. Only the proposed method and CSK algorithm can effectively track targets in thick cloud image sequences.

When the target is completely occluded, as shown in the third line of Figure 8, the features of small and medium-sized targets in frame 189 are almost lost, and the position of small targets can only be estimated by the position of the targets in the previous frame. At this time, the CSK algorithm, which performed well before, has lost the target, while KCF algorithm and DFT algorithm have failed before entering the cloud. Only the method in this paper can deal with the target tracking problem under complete occlusion This is because Kalman filter is added to predict the target position.

For the image sequence in dark background, as shown in the fourth line of Fig. 8, the background clutter of small target is very large, and the target is far away from the ground. KCF algorithm and DFT lost the target at the beginning of tracking. The only way we can track our targets is through CSK. However, it is a pity that the problem of tracking drift occurs in the tracking process of CSK algorithm. For example, the tracking drift occurs in the sixth frame of sequence 3. Our method uses APCE to modify the tracking position continuously to solve the problem of tracking drift.

In the image sequence 5, the fifth line of Fig. 8 is shown. The target is the launched rocket. Due to the rapid movement, CSK algorithm, DFT algorithm and CXT algorithm can not adapt to the high-speed movement of the target, so the tracking failure occurs at frame 145. The KCF algorithm and the proposed method can always track the target, but the proposed method has not occurred tracking drift, and the KCF algorithm has appeared target tracking drift since the 145th frame.

In the image sequence 8, as shown in the sixth line of Fig. 6, it is the sea sky background with low resolution. The CXT algorithm fails to track the small target at the sea sky boundary because of the low contrast between the target and the sea level. The proposed method and the other three methods can track small targets. Our method has the highest overlapping rate of bounding box and ground-truth, because our method uses APCE as the threshold to correct when updating the template.

The last image sequence is a low contrast cloud sky scene, as shown in line 8 of Figure 7. CXT algorithm has the worst performance as before. KCF algorithm, CSK algorithm and DFT algorithm all produce tracking drift, which is because they do not consider the complexity of small target tracking scene, so it is not working to apply the traditional correlation filtering algorithm directly in the field of small target tracking.

- Objective experimental analysis

We used vot2016 benchmark evaluation standard to analyze the experiment. The evaluation indexes are two indicators, namely accuracy rate and success rate. Firstly, the success rate is defined. Here, the coincidence rate score is defined as the coincidence rate of the bounding box obtained by the ground truth and tracking algorithm. When the success rate of a frame is greater than the set threshold, the tracking is successful. The success rate is defined as formula 13, where n is the number of successful tracking and N is the total number of frames.

$$\text{Success} = \frac{n}{N} \quad (13)$$

Next, we introduce the accuracy rate, which is the percentage of video whose distance between the center of the bounding box and the center of the ground truth is less than a given threshold.

The experimental results are shown in Figure 7. The first line of Fig. 7 shows the accuracy map under different tracking backgrounds, and from left to right are the overall accuracy map, the tracking accuracy map under the background clutter condition, the tracking accuracy map under the condition of fast moving small target and the tracking accuracy map under occlusion condition. The second line shows the success rate graph under different tracking backgrounds. From left to right, it is the success rate graph under all sequences, the success rate graph under background clutter, the success rate graph under fast moving target and the success rate graph under occlusion condition.

In terms of accuracy, the proposed method has surpassed the traditional KCF algorithm in all aspects. In terms of success rate, the proposed method is also the best one, reaching a success rate of 0.452 when the overlap rate threshold is 0.5.

In terms of experimental speed, CSK algorithm and KCF algorithm can quickly track the target under the condition of simple tracking scene, and the operation speed reaches 662.79fps and 328.60fps. The proposed method sacrifices the speed and accurately tracks the target, with the speed of only 35.80fps. However, it is still higher than CXT algorithm and DFT algorithm. See the table 3 for detailed speed.

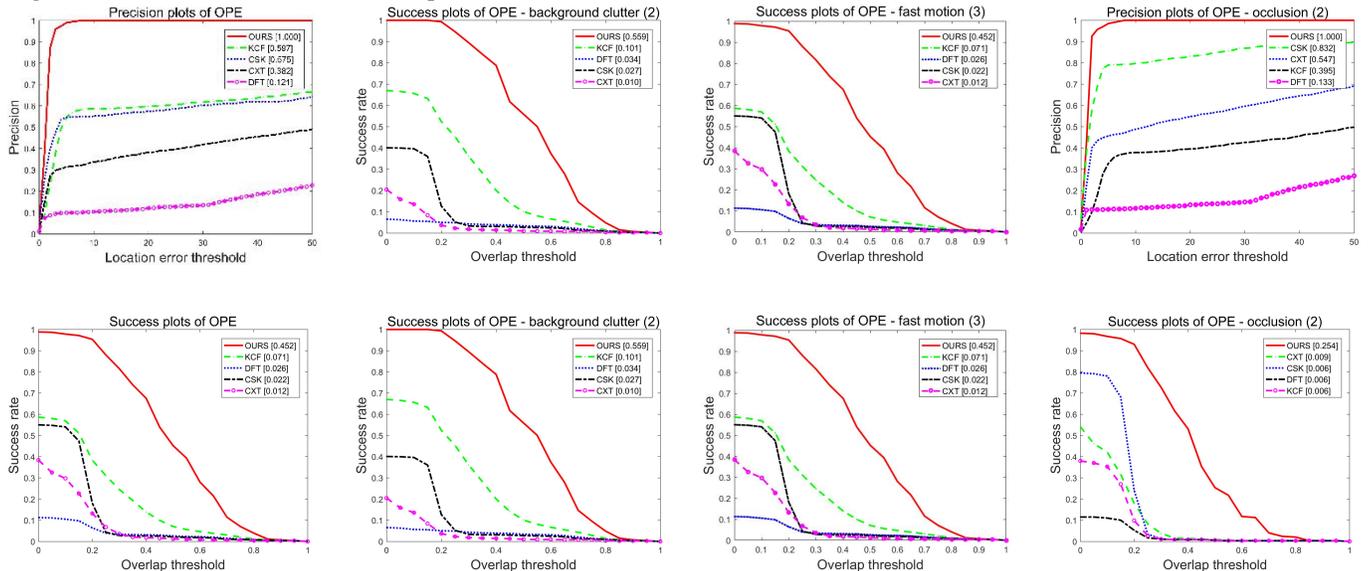


Figure 7. Accuracy map and success rate graph of five methods in different scene sequences.

Table 3 Mean FPS comprehensive evaluation

Algorithms	CSK	CXT	DFT	KCF	OURS
Mean FPS	328.80	24.20	21.53	662.79	33.35

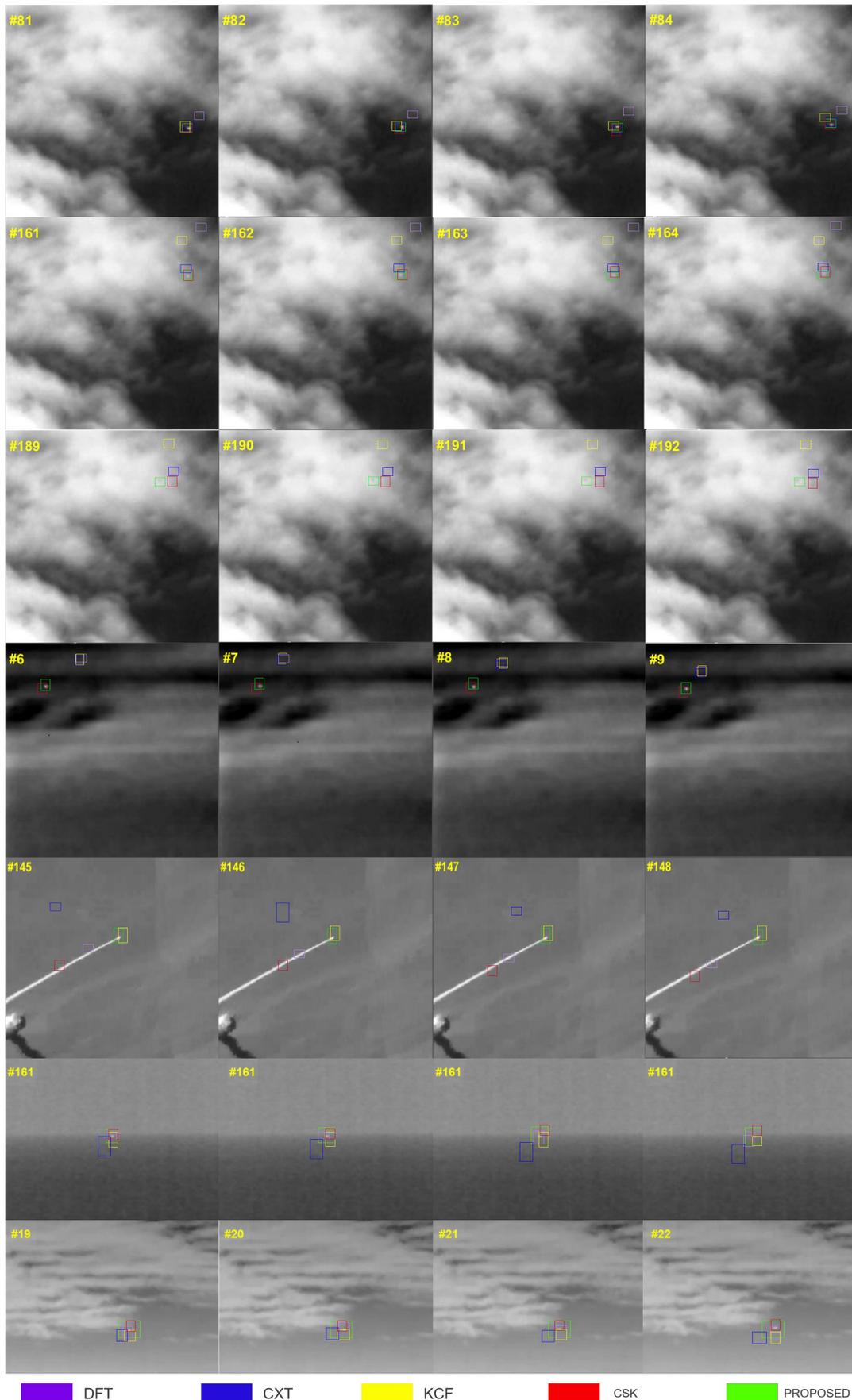


Figure 8 The experimental results are intuitionistic. The purple box is the DFT algorithm result graph, the blue box is the CXT algorithm result, the yellow box is the KCF algorithm tracking result, the red box is the CSK algorithm tracking result, our algorithm is represented by the green box.

Conclusion

In this paper, an algorithm for long-term tracking of small infrared targets is proposed. The problem of target occlusion and drift can be solved by using multi feature reconstruction measurement framework. Finally, four groups of image sequences and four comparison methods are used for experimental analysis. It is concluded that the performance of the proposed algorithm is good. Reducing the time complexity of the algorithm and applying our algorithm to multi-target tracking are the future work.

References

1. P. Du and A. Hamdulla, "Infrared Moving Small-Target Detection Using Spatial–Temporal Local Difference Measure," in *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 10, pp. 1817-1821, Oct. 2020.
2. Y. U. Sinn, K. M. Hopkinson, B. J. Borghetti and B. J. Steward, "IR Small Target Detection And Prediction With ANNs Trained Using ASSET," 2019 IEEE Aerospace Conference, Big Sky, MT, USA, 2019, pp. 1-11.
3. P. Du and A. Hamdulla, "Infrared Moving Small-Target Detection Using Spatial–Temporal Local Difference Measure," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 10, pp. 1817-1821, Oct. 2020.
4. W. Wu and Askar Hamdulla, "Space-time-based Detection of Infrared Small Moving Target", *Journal of Applied Science and Engineering*, vol. 23, no. 3, pp. 497-507, may. 2019.
5. Kun Qian, Huixin Zhou, Shenghui Rong, "Infrared dim-small target tracking via singular value decomposition and improved Kernelized correlation filter," *Infrared Physics & Technology*, vol.82, pp.18-27. February 2017.
6. Z Li, J Li, F Ge, "Dim moving target tracking algorithm based on particle discriminative sparse representation, " *Infrared Physics & Technology* , v1o, 75, pp 100-106. January 2016.
7. M Shi, B Wang, and Y Cao. "A Kernel Correlation Filter Tracker with Fast Scale Prediction, "In *Proceedings of the 2020 12th International Conference on Machine Learning and Computing (ICMLC 2020)*. ACM, NY, USA, PP 574–579.
8. W Hao, "A gradient descent method for solving a system of nonlinear equations, " *Applied Mathematics Letters*, vol. 112, 2021.
9. Y He, M Li, J Zhang, J Yao, "Infrared target tracking via weighted correlation filter, " *Infrared Physics & Technology*, vol.73, PP. 103-114. September 2015.
10. B Bai, B Zhong, G Ouyang, "Kernel correlation filters for visual tracking with adaptive fusion of heterogeneous cues, " *Neurocomputing*, vol.286, PP. 109-120. January 2018.
11. M. Wang, Y. Liu and Z. Huang, "Large Margin Object Tracking with Circulant Feature Maps," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 4800-4808.
12. F. Henriques, R. Caseiro, P. Martins, and J. Batista. Exploiting the Circulant Structure of Tracking-by-Detection with Kernels. In *ECCV*, 2012.
13. T. B. Dinh, N. Vo, and G. Medioni. Context Tracker: Exploring supporters and distracters in unconstrained environments. In *CVPR*, 2011.
14. T. B. Dinh, N. Vo, and G. Medioni. Context Tracker: Exploring supporters and distracters in unconstrained environments. In *CVPR*, 2011.
15. M. Wang, Y. Liu, and Z. Huang, "Large Margin Object Tracking with Circulant Feature Maps," *IEEE Computer Vision & Pattern Recognition*, pp. 4800-4808, 2017.

Figures

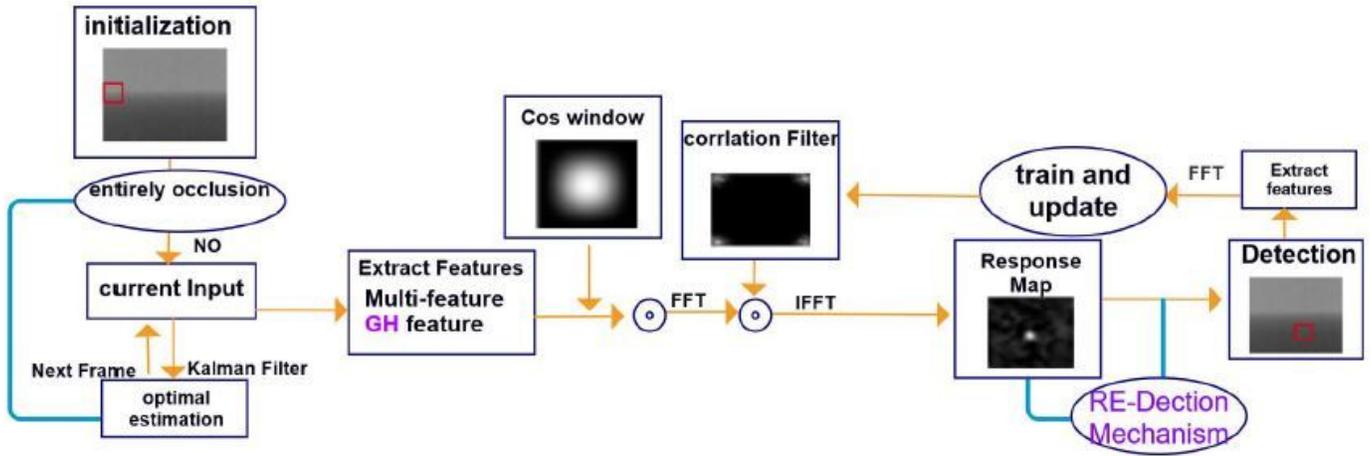


Figure 1

algorithm flow chart

$$F = \lambda F_{gary} + \alpha F_{HOG}$$

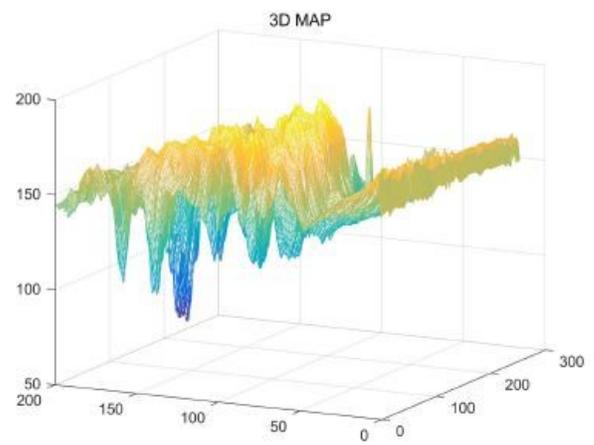
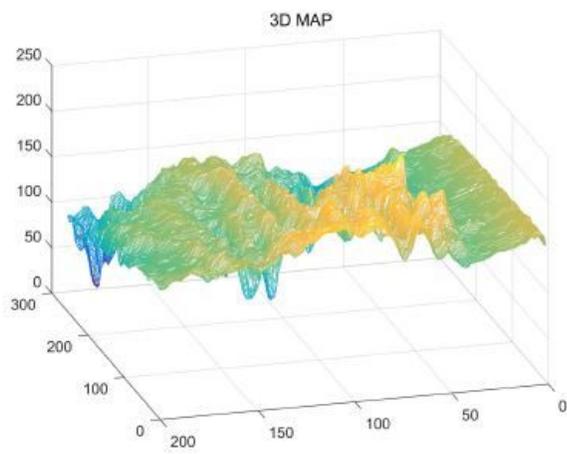


Figure 2

3D image under occlusion condition and 3D image without occlusion

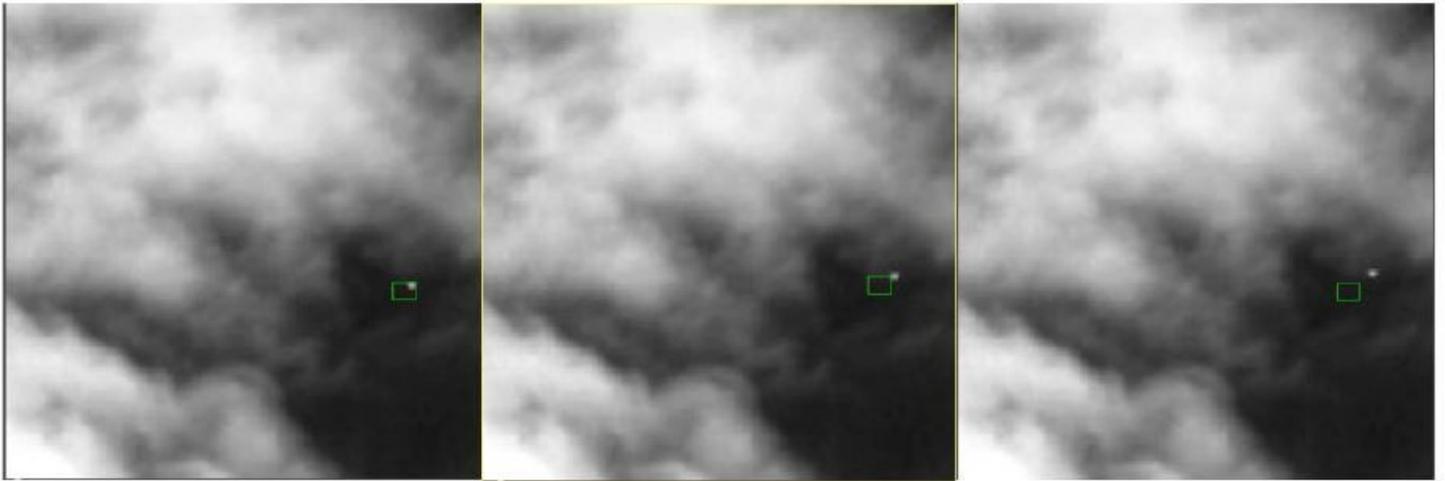


Figure 3

KCF algorithm fails to track when the target enters the cloud. (The green box represents the tracking result.)

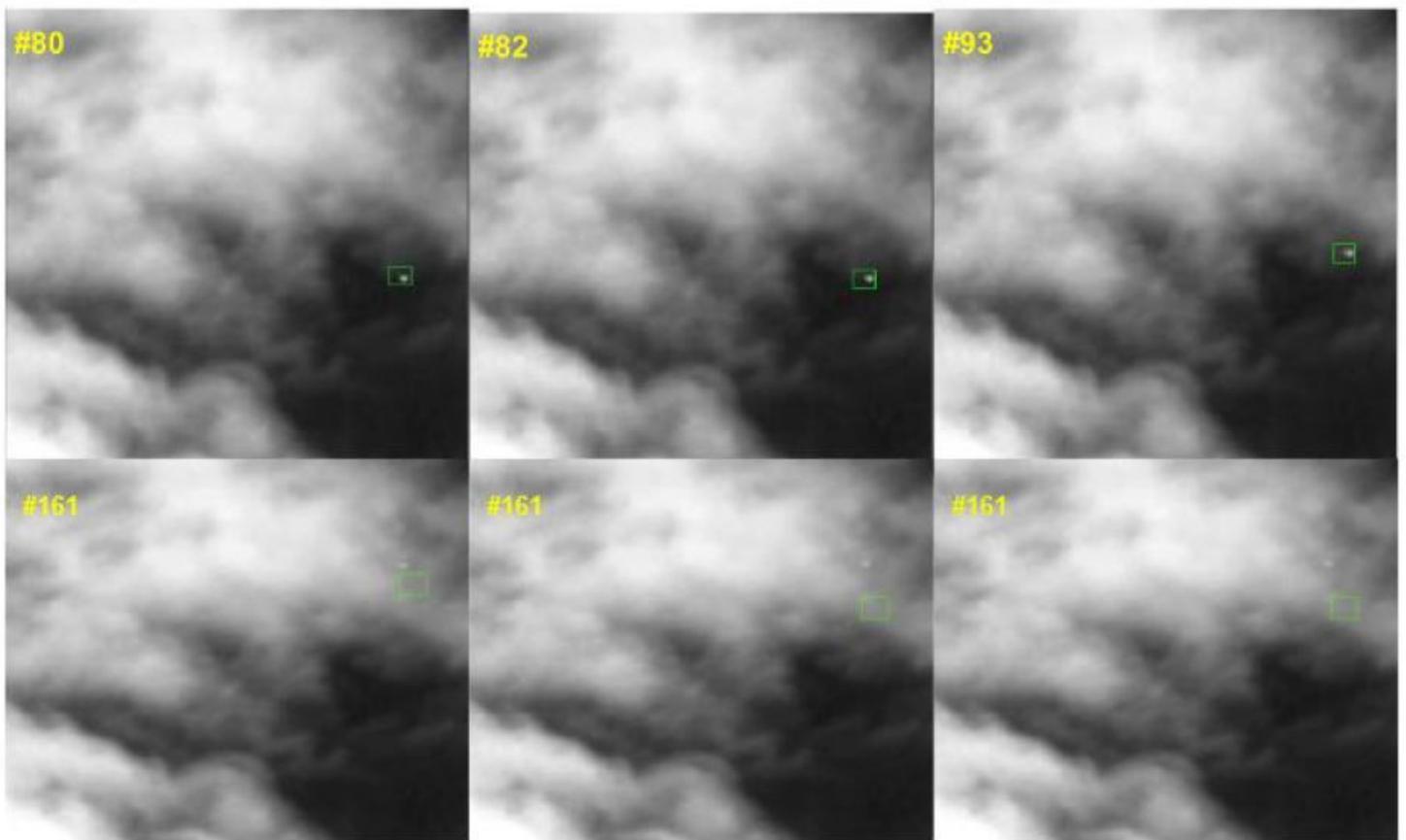


Figure 4

Tracking result map after adding gray feature and hog multi feature function.

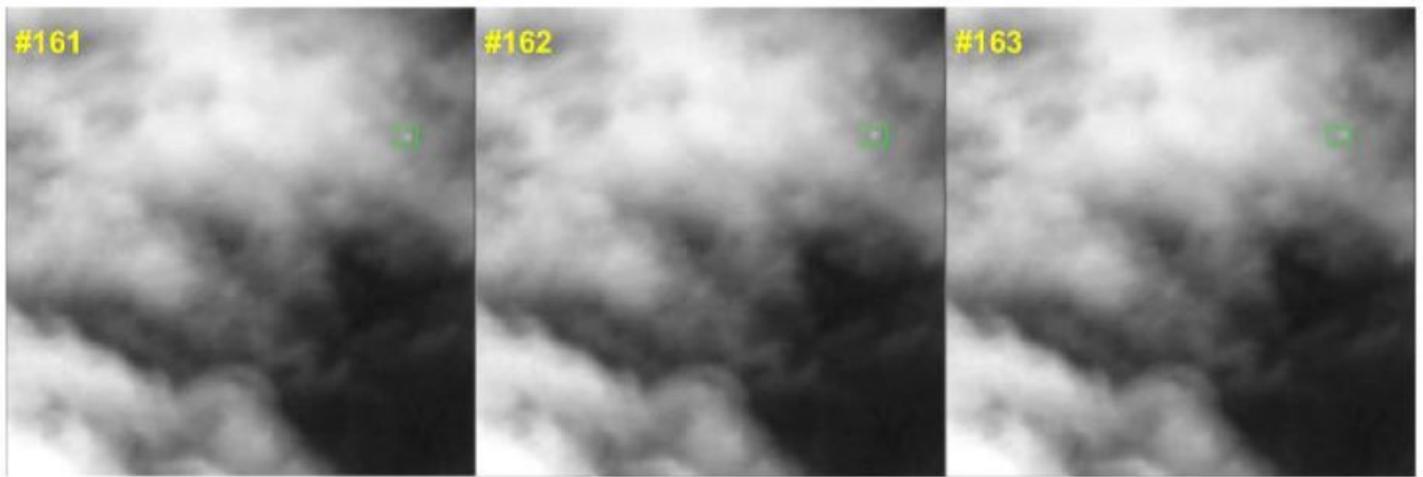


Figure 5

Tracking results of algorithm with new multi feature function

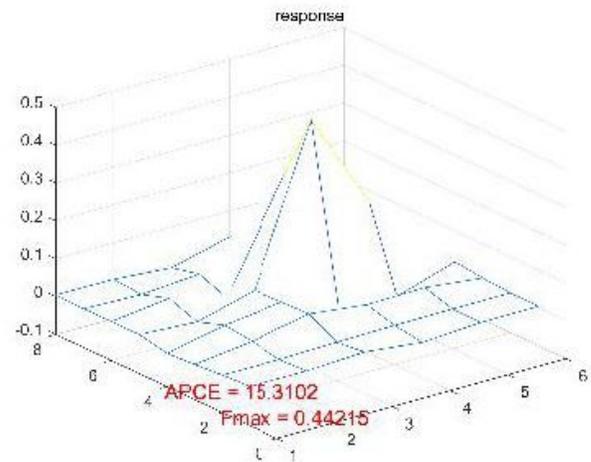
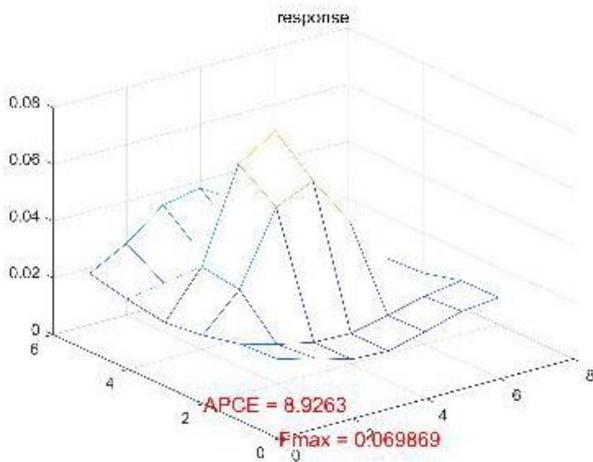


Figure 6

The difference map of APCE between occlusion and non occlusion, The first behavior is the original image with occlusion and the original image without occlusion, and the second behavior is the corresponding APCE.

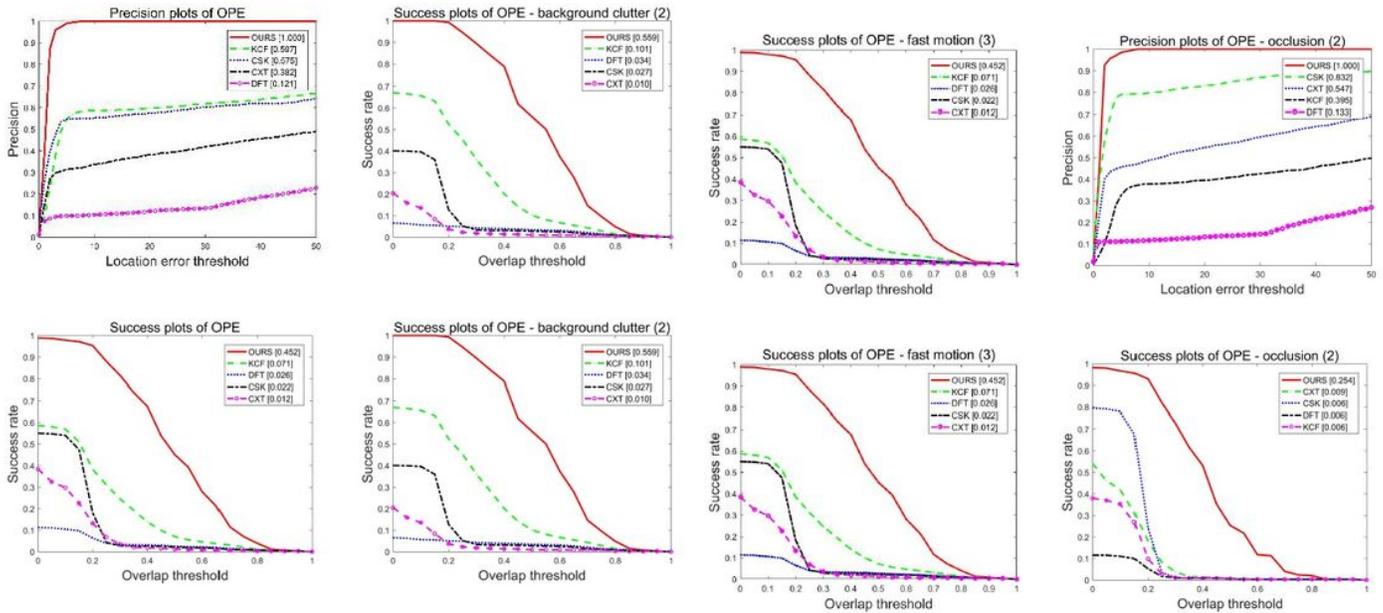


Figure 7

Accuracy map and success rate graph of five methods in different scene sequences.

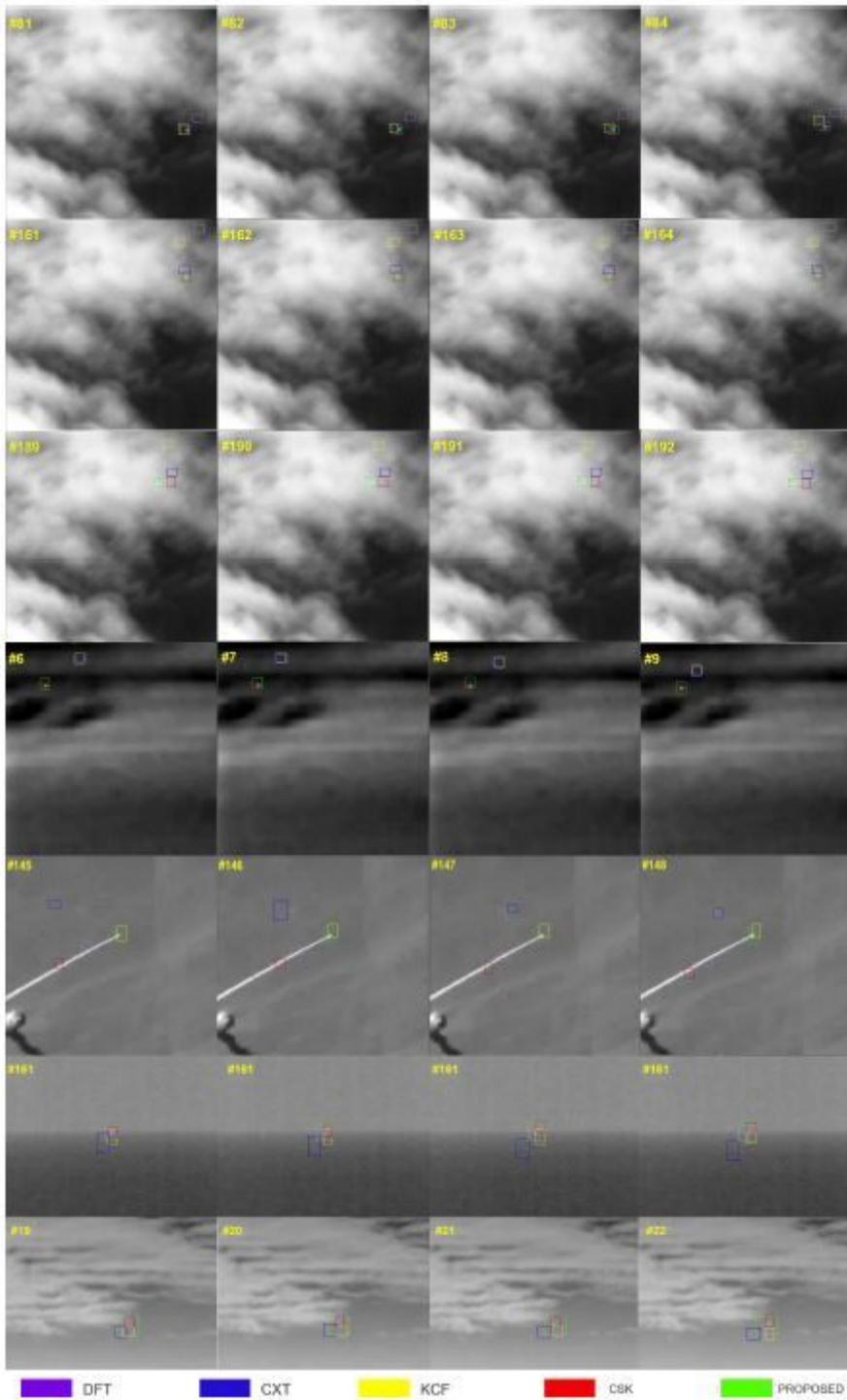


Figure 8

The experimental results are intuitionistic. The purple box is the DFT algorithm result graph, the blue box is the CXT algorithm result, the yellow box is the KCF algorithm tracking result, the red box is the CSK algorithm tracking result, our algorithm is represented by the green box.